7-10-1958

The Use of Objective Tests in Attaining Industrial Arts Objectives

Earl Shedlebower

THE USE OF OBJECTIVE TESTS IN ATTAINING INDUSTRIAL ARTS OBJECTIVES

EARL SHEDLEBOWER

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science in Education
(Plan B)

Approved

July 10 1958 [redacted]
Dr. R. H. Landis
Advisor

July 10 1958 Walter A. Klehm
Dr. W.A. Klehm
Advisor

Eastern Illinois University
Charleston, Illinois
Summer, 1958

Table of Contents

THE USE OF OBJECTIVE TESTS IN ATTAINING INDUSTRIAL ARTS OBJECTIVES

Introduction

Throughout our lives we are constantly confronted with the problem of measurement and evaluation. This is true to the extent that we accept it as a common fact in our every day living. The clock measures time; the automobile odometer measures distance; and the speedometer measures speed.

Other types of measurement determine the value of one item over another. W hen comparing for exactness, the process is commonly referred to as testing. These types of testing are also taken for granted. A sampling of the clothes we buy, the automobiles we ride in, and the household items we use were tested before the products were released to the public.

Another area of testing is that of giving tests to individuals. Achievement tests measure the achievement of the student, aptitude tests measure the potential of the student, and interest inventories determine a person's likes or dislikes.

The purpose of this paper is based on one specific area of testing or the use of objective tests in attaining industrial arts objectives.

It is not the intention of the writer to determine the actual worth of the objective test as compared with other tests. However, since the advantages, limitations, procedures of construction, validity, and uses and values of the objective test will be discussed some comparison is necessary,

Advantages of the Objective Test

The objective test has advantages which are peculiar to it alone.

T he economy involved in giving an objective test demands that no unnecessary writing be done. W hen the discussion-type of test is given, the time available demands that fewer questions be used. Many questions would be used if the test were an objective type. The reliability of a ten question test can hardly be compared with a one-hundred question test.[1] When comparing the outcomes of these two types of tests, it must be understood that the items have been carefully selected.

When using a discussion type of test, the time involved to give enough of a sampling for the test to be of value would be lengthy. However, the time required to accomplish the same results with an objective test would be short.[2]

---

1. H. D. Rinsland, Constructing Tests and Grading, (New York, Prentice-Hall, Inc., 1937), p. 278.

2. D. G. Patterson, Preparation and Use of New-Type Examinations, (Chicago, World Book Co., 1925), p. 11.

To illustrate this point, comparisons have been made to determine the approximate length of time required for the discussion test to equal the reliability of the objective test.[3] It was found that approximately three weeks of discussion type testing was necessary to equal approximately two hours of objective type testing.

In some instances, this point may be without substance. When comparing the time spent in making up tests, with the relative reliability of each, there is little difference.[4]

When comparing the reliability of a test per unit of working time, it must be taken into consideration the exact comparison that is intended. If the reliability of the test is based on the number of responses, which gives a broader basis for determining student achievement, then the test is quite reliable.[5]

Another advantage of the objective test is that the student is required to give a direct answer. This answer may be given in many forms. The type of question will usually determine the form of answer expected. When using this type of test, the student must know exact facts and be able to relate them as required.

---

3. G. M. Ruch, The Objective or New-Type Examination, (Chicago, Scott, Foresman and Co., 1929), p. 116.

4. Ibid., p. 116.

5. H. A. Greene and A. N. Jorgensen, The Use and Interpretation of High School Tests, (New York, Longmans, Green and Co., 1936), p. 110.

This type of testing will discourage bluffing. In subjective testing, a student may bluff his way through by writing much about nothing. He often hopes that quantity rather than quality will suffice. Other students often rely on the good nature of the instructor, hoping that they will not receive an absolute zero because they tried. It is not correct to say that bluffing has been entirely eliminated by the use of the objective test, but it has been reduced to a point where it is not objectionable.[6]

Objective scoring has an advantage over subjective scoring. When the scoring of a test is left open to the opinion of the instructor, the outcome is not always the expected. When tests are scored objectively, the answer can be classed as right or wrong. In this manner, the student is measured directly; either he knows the answer or he doesn't.

However, when the personal opinion of the instructor enters into the scoring, his whims, temperament, and feelings may influence the score given the student. Actually, objectivity is an essential of objective measurement. The educational process may be highly subjective, but measurement implies accuracy.[7] Therefore, it would seem that if the student were to receive a fair evaluation of his responses, then this source of inaccuracy should be eliminated to a desirable extent.

---

6. Ruch, op. cit., p. 109.

7. Ruch, op. cit., p. 113.

Extensive sampling is possible with the use of objective tests. The reliability of a test may be based on two factors, objectivity and extent of sampling. The former has been discussed previously in this paper. Unreliability due to sampling cannot be completely eliminated since measurement is a sampling, never a complete coverage of the subject.[8]

The traditional type of test is based on extensive coverage of a small number of items, while the objective type test relies on the sampling of many topics with less emphasis being placed on each one.

The extent of sampling has bearing on the reliability and validity of a test. It is important that the sampling be extensive enough for the test to be reliable. However, as insufficient sampling can cause an unreliable test, an over-sampling may cause the test to become an end in itself rather than a means to an end.[9] Therefore, a correct amount of sampling is necessary to secure an accurate measure of each pupil.

Another advantage, but perhaps of less importance, is better control of the test by the instructor. Pupils who are in doubt on the answer to a discussion question, may often get by on the pretense that they misunderstood the question. By the use of the objective type test, the instructor requires the student to react in the intended manner.

---

8. Ibid., p. 115.

9. Rinsland, op. cit., p. 301.

This should not be objectionable since the teacher exercises the same right when controlling student reactions during class instruction.[10]

Limitations of the Objective Test

The limitations of an objective test must be taken into consideration for an unbiased discussion. As the advantages of an objective test point out its good points, so the limitations place an emphasis on its weak points.

One such limitation which has given rise to much discussion is that objective tests tend to measure memory only.[11] The two opposing points of view on this statement are concerned with the thought versus the memory question.

The thought question which requires opinion and relating of ideas on the part of the student is misused many times. Those who advocate its use often ask pupils to list items which are credited to memory. Contrasting with the thought question is the memory question. This type of question concerns itself with a response formulated from the student's memory.

---

10. Ruch, op. cit., p. 120.

11. Ibid., p. 120.

The difference between the supposedly poor memory question and the thought question is not in the questions but in the background of the student.[12] One student may answer a question by thought and reason while another student may have the answer committed to memory.

An objective test may be open to guessing and chance. Under these conditions, the test is not considered reliable.[13] The extent to which this is true will depend on the characteristics of the test. Under certain conditions, the answering of test items is based on some physical form. Many students consider that the mere flipping of a coin should determine the response to a question.

Many of the chance and guessing responses are not chances. Usually the student makes a response with sufficient knowledge causing him to guess rather than to rely on pure chance.[14]

Perhaps the element of guessing and chance can be eliminated by giving students instructions against guessing. This is a point of discussion which has never been completely settled.[15] Many instructors advise students to guess, thereby increasing the possibility of a better score.

---

12. Ruch, op. cit., p. 120.

13. Ruch, op. cit., p. 356.

14. L. V. Newkirk and H. A. Greene, Tests and Measurements in Industrial Education, (New York, John Wiley and Sons, Inc., 1935), p. 124.

15. C. W. Odell, Educational Measurement in High School, (New York, The Century Co., 1930) p. 475.

This may be a good practice when only one point is subtracted for a wrong answer. When a two-off-for-one-wrong answer system is used, the student is penalized twice, lessening his chances of receiving a better score by guessing.

When using objective tests, the chance for self-expression by the student is lessened. This may not be objectionable limitation since the objective test is primarily used as a measurement for accuracy.[16]

The fact remains that there may be times when the student is capable of making an acceptable response by self-expression. Proper provisions should be made for this type of expression. This may be accomplished by using a modified objective type of test.

Many objective type tests depend more on selection among alternatives than on spontaneous recall. The traditional type test relies mostly on the latter type of response.[17] This may be a direct limitation, for a spontaneous recall response demands a vast amount of stored information. The selection of alternatives on the other hand requires that only the right information be on hand when a choice is given.

An additional limitation which is of lesser importance than those previously discussed is that the objective type test usually requires some form of duplicating.

---

16. Rinsland, op. cit., p. 304.

17. Ruch, op. cit., p. 128.

The problem of duplication is of very little importance in view of the fact that the cost of master copy sheets, paper and duplicating equipment is small in comparison with other school expenses.[18]

Procedures of Construction of the Objective Test

The procedure involved in constructing a reliable and valid test may involve many hours of work carried out according to a definite scheme. This does not mean merely the putting down of questions as they come to mind, but requires the following of a systematic step by step procedure. If the instructor works in this manner, he will be assured of success. This type of procedure will prove valuable when writing short as well as long tests.[19]

The conditions under which the test is to be given should be taken into consideration early in the test construction. The instructor should determine the general aims which he is expect- to test. These may be the general aims of the program or of the course. To supplement this list, the specific aims should be determined.[20]

---

18. Ruch, op. cit., p. 126.

19. W. J. Micheels and M. R. Karnes, Measuring Educational Achievement, (New York, McGraw-Hill Book Company, Inc., 1950), p. 126.

20. C. C. Ross, Measurement in Todays Schools, (New York, Prentice-Hall, Inc., 1947), p. 104.

Such a list of aims will enable the instructor to be more accurate when constructing the test items.

Upon completion of the list of general and specific aims, it is necessary to analyze them for the outcomes which are expected. Analyzing the aims may take considerable time, but it is necessary because an analysis will reveal much valuable and lasting information.[21] The instructor will have several points which he expects the test to bring out. If a majority of these appear in the analysis, the test is of value. However, if many of the expected outcomes are not forthcoming, it may be necessary to carefully re-analyze the test to determine the reason or reasons for this failure.

A general guide is needed about which to build the test. Such a guide is called a "table of specifications".[22] This guide is necessary to insure against omitting important items and to assure a balance in the sampling. The guide will also serve as a working plan when actual construction of the items begin.

After a working plan has been completed, the drafting of the items is begun. At first, tentative test items should be written which cover the main points of each topic or aims of the course. The type of item which will best fit the point should be a determining factor in the construction of the test items.[23]

---

21. Micheels and Karnes, op. cit., p. 127.

22. Ruch, op. cit., p. 150.

23. Micheels and Karnes, op. cit., p. 127.

A point may be well presented in one form but may be stated much better in another. This will tend to mix the various forms of test items. This is of little importance since they may be easily separated later.

The next step in test construction is to arrange the test items into groups of the same type. Different items of the same type may be arranged so that related items are together. The main reason for doing this is to save time on the part of the student. Arranging the items in order of difficulty is best done after the test has been given.[24]

After completing the steps listed so far, the test is then ready to be put into a form for administering. At this time, clear and specific directions should be written. Different directions are needed for each type of question. This is necessary in order to inform the student exactly what to do and how to do it.

When these steps have been completed, the test should be assembled and proof-read.

To further improve the test, it would be wise to submit the test to another instructor for comments and suggestions for improvement. The test should then be revised, including the necessary items or statements and eliminating those having no value. This should be done by keeping in mind the comments and suggestions of the other instructor.

---

24. Micheels and Karnes, op. cit., p. 198.

A key or score sheet should be made out before the scoring is begun.[25] Many items, a list of the correct answers will suffice. The key or score sheet will aid the instructor by saving valuable time. It is very time-consuming to read every question and the student's response. Often, the key or score sheet is arranged so it may be placed close to the answers, thereby making a rapid comparison possible.

The rules for scoring should also be determined before scoring is started. This will assure each student a fair score. it is unwise to give a student a grade unless it has been done by a set of rules.[26] In this manner, there is no chance for a controversy to arise because one student received a better rating than another.

After the revisions and the key and score sheet are made, the test is ready to be administered to the students. After the test has been given, it should be analyzed and revised for improvement. A constant revision is necessary to provide an improved and valid test.

A method which will be of value in building better tests is the compilation of a card file of various test items. A card file of test items is a handy reference and storage place for effective items, and the file may be increased as new items are introduced.

---

25. Ross, op. cit., p. 123.

26. Ruch, op. cit., p. 184.

Usually there are certain aims which will change little. Therefore, a large number of items may be formulated around this assumption.[27] Each aim should be cataloged, and the specific items relating to that aim should be placed under that heading. Constant improvements will be necessary to keep the file up-to-date.

Validity of the Objective Test

The validity of a test may be determined by the accuracy with which it measures what it is supposed to measure.

The validity of a test depends upon the care taken to include items which are of value and to eliminate those which have no value.[28] A test which measures abilities in an accurate fashion has validity and a test which measures abilities for which it was not intended has little validity.[29]

To avoid a misunderstanding at this point, it might be wise to point out that validity is a much broader term than reliability. A test may be reliable and not be exceedingly valid. However, a test which is valid would, to a great degree, be a reliable test.[30]

---

27. Micheels and Karnes, op. cit., p. 129.
28. Ruch, op. cit., p. 28.
29. Micheels and Karnes, op. cit., p. 104.
30. Micheels and Karnes, op. cit., p. 112.

There are several principal methods of validating tests. Expert opinion in the form of courses of study, textbooks, judgments of persons well versed on validating tests, and the recommendations of committees and organizations have proved to be reliable methods of validating a test.[31] Other methods of validating a test are experimental in nature. One method is to study the errors which occur most frequently. A second method is to correlate the test against another accepted valid test. A third is to write down good items throughout the year; this eliminates reliance on memory when it is time for a test and the possibility of including items which are of little value. A fourth method is to have a fellow teacher look over the work. It is understood that the second teacher may not be an expert but he may suggest several good ideas for changing or improving the test.[32]

There is no hard and fast rule which will insure the construction of a valid test. A test which is valid for one situation is not necessary valid for another.[33] To determine the true validity of a test, it should be evaluated against the situation for which it was intended.

---

31. Ruch, op. cit., p. 29.

32. Odell, op. cit., p. 54.

33. Wicheels and Karnes, op. cit., p. 105.

The following factors in order of importance influencing the validity of a test in a given situation are: the characteristics of the students involved; the time at which the test is given; the type of test, whether it be standardized or teacher-made.[34]

Students affect the validity of a test to a great extent. The student may be a poor reader, his power of memory may be poor, or his reaction time may be slow.

The use of the so-called standardized test may prove reliable, but the validity of such a test under certain circumstances is low. The standardized test is reliable in that it measures consistently that which it measures. However, the validity may be challenged on the ground that it does not test what it is supposed to test.[35]

Not only is a test determined valid as a whole, but each item may be termed valid or invalid.[36] In general, the validity of the items will determine the validity of the test as a whole. A valid test will be an assurance that the individual items do what they are supposed to do.

It would be of more use if the individual items were given attention first to assure valid test items. However, the overall validity of a test is much more than the addition of independently valid test item.[37]

---

34. Micheels and Karnes, op. cit., p. 112.

35. Micheels and Karnes, op. cit., p. 106.

36. Ross, op. cit., p. 78.

37. Micheels and Karnes, op. cit., p. 103.

The construction of valid test items is determined by several principles. These are, wording of the items, difficulty of items and use of catch questions. The wording of a test item should be such that most students can readily understand its meaning. Improper wording will tend to lessen the validity of a test item.

The relative difficulty of a test question will also influence its validity. The instructor should eliminate those questions which all students answer correctly and those which all students fail to answer correctly. This does not mean that these questions have no validity. They may be valid under other conditions.

The use of catch questions are usually unsatisfactory. They tend to lead the student in the wrong direction. When this occurs, the items lose their validity.[38]

Although validity may be kept in mind during the construction of the test, its validity cannot be proved until it has been administered to the students.[39]

Uses and Relative Values in Attaining Industrial Arts Objectives

After the true validity of a test is determined, statements can be made as to its probable use and relative value in attaining industrial arts objectives.

---

38. Odell, op. cit., p. 478.

39. Odell, op. cit., p. 59.

In recent years the objective test has been accepted as being related to instructional material.[40] Therefore, by using the test as he teaches, the instructor is better able to discover how to chart his course.

The necessary conditions which produce this type of test require much of the teacher's time. However, he should not be expected to rely upon observation to determine differences of the students.

The objective test is of value when used to measure the achievement of the student.[41] This is best accomplished when the test items cover the material taught. If this is kept in mind, then an accurate measurement is obtained. This type of testing may be varied by testing for achievement at the beginning and then at the end of the class.

The use of the objective test will also prove to be of value as a motivating device.[42] A problem does arise when the test is used with this value in mind. If the construction is poor, it will tend or permit poor responses by the student.

Contrary to this, the test which is so constructed that it promotes good work habits on the part of the student will possess the greatest motivating power. Therefore, it will be of value.

The test will be of more value if it is used as a source of information which will make the students aware of their accomplishments and limitations.[43]

---

40. Newkirk and Greene, op. cit., p. 18.
41. Newkirk and Greene, op. cit., p. 19.
42. Rinsland, op. cit., p. 12.
43. Green and Jorgensen, op. cit., p. 77.

When the objective test is used in this manner, it will increase the efficiency of instruction.

The efficiency of instruction as a value of the objective test is very important.[44] A problem arises when low scores show up on a test. This problem is much increased when the reason for the low scores is not known. This problem is greatly lessened when the results of the test are analyzed and the reasons for low scores are discovered. Perhaps the items were poorly worded or the directions were insufficient. Another possibility is that the students were at fault.

If the latter is true, the use of the test may possess another value. This value is class diagnosis. A better situation in class is afforded if the instructor is aware of the strengths and weaknesses of the students. Therefore, he is better able to strengthen the weak points of his instruction.

The value of class diagnosis is valuable to the student as well as to the instructor.[45] It is valuable to the students in that those who show prior knowledge of the material need not be required to repeat that part of the subject matter. This is also true on the part of the instructor. If he is aware of a student's ability, he will be less likely to waste time giving that student repeated information.

---

44. Michaels and Karnes, op. cit., p. 86.

45. Newkirk and Greene, op. cit. p. 24.

Another value of an objective test which is often stressed above other values, is that it serves as a basis for a grade.[46] Often a student is given a grade on a test score alone. The chief danger in this is that the true worth of a student cannot be decided by a test score. However, the use of several such test scores will provide a partial basis on which the student's accomplishments may be evaluated objectively.

---

46. Green and Jorgensen, op. cit., p. 46.

Summary

In the area of written measurement and evaluation in industrial arts, two forms of test items are generally recognized. They are the "subjective", in which the opinion or judgement of the instructor may largely influence the rating of the items; and the "objective", in which the personal bias of the instructor is eliminated to a large extent. Of the two, the latter type is generally regarded as being more valuable in securing an accurate and fair judgement of achievement.

The advantages of the objective type test are:

1. It is economical to administer.
2. It is quite reliable per unit of working time.
3. It discourages "bluffing" on the part of the student.
4. It is scored objectively, thus eliminating the judgement of the instructor.
5. Extensive sampling is possible giving an accurate measure of each student.
6. The student is required to give a positive answer.
7. The instructor has better control of the test by requiring the student to react in a certain manner.

The limitations of the objective type test are:

1. The test tends to measure memory of facts only.
2. The test is open to guessing and chance based on some physical form.

3. The possibility for self-expression on the part of the student is lessened.
4. The answers to the test depend upon the selection among given alternatives.
5. The test requires some form of duplicating in order that each student will possess a copy of the test.

The procedures required to construct a valid and reliable objective type test are:

1. To determine the physical conditions under which the test is to be given.
2. To list and analyze the aims of the course for the expected outcomes.
3. To draw up a "table of specifications" around which to build the test.
4. To begin drafting the items and arranging them into groups of the same type.
5. To put the test into a form for administering to the students.
6. To submit the test to another instructor for comments.
7. To revise the test keeping in mind the comments of the other instructor.
8. To write clear and specific directions to the students for working out the test.
9. To make out a score sheet and determine the rules for scoring to assure each student a fair evaluation.

The extent to which a test measures what it should measure depends upon the following factors:

1. Only items which measure what they should measure should be included in the test.
2. Personal differences of the students taking the test may influence the degree to which the test measures what it should measure.
3. The time and type of test, whether formal or teacher made, may affect the degree to which a test measures what it should measure.
4. The relative difficulty and the use of catch questions usually affect the degree to which a test measures what it should measure.
5. The test should be evaluated against the situation for which it was intended, to determine the degree to which the test measures what it should measure.

The uses and relative values of the objective type test are:

1. It enables the instructor to discover how to chart his course.
2. It measures the achievement of the student over the material covered.
3. It encourages good work on the part of the student.
4. It provides a basis on which the instructor may strengthen the weak points of his instruction.
5. It provides the instructor information about the strengths and weaknesses of his students.
6. It provides a sounder base for the issuing of individual student grades.

References

1. Greene, H. A. and Jorgensen, A. N., The Use and Interpretation of High School Tests, (New York, Longmans, Green and Co., 1938), pp 583.

2. Micheels, W. J. and Karnes, M. R., Measuring Educational Achievement, (New York, McGraw-Hill Book Company, Inc., 1950), pp 488.

3. Newkirk, L. V. and Greene, H. A., Tests and Measurements in Industrial Education, (New York, John Wiley and Sons, Inc., 1935), pp 241.

4. Odell, C. W., Educational Measurement in High School, (New York The Century Co., 1930), pp 315.

5. Patterson, D. G., Preparation and Use of New Type Examinations, (Chicago, World Book Company, 1926), pp 78.

6. Remmers, H. H. and Gage, N. L., Educational Measurement and Evaluation, (New York, Harper and Brothers, 1943), pp 560.

7. Rinsland, H. D., Constructing Tests and Grading, (New York, Prentice-Hall, Inc., 1937), pp 314.

8. Ross, C. C., Measurement in Todays Schools, (New York, Prentice-Hall, Inc., 1947), pp 533.

9. Ruch, G. M., The Objective or New-Type Examination, (Chicago, Scott, Foresman and Company, 1929), pp 445.